J. Math. and Its Appl.
ISSN: 1829-605X
Vol. 12, No. 1, Mei 2015, 13-22

SIFAT-SIFAT GENERALISASI DISTRIBUSI BINOMIAL YANG BERTIPE COM-POISSON

Farida Agustini Widjajati[1], Marselly Dian Saputri[2], Nur Asiyah[3]
[1,2,3]Jurusan Matematika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111
[1]agustini.farida54@gmail.com

Abstrak

Distribusi Binomial dan Poisson digunakan untuk menganalisis data diskrit. Karena distribusi Poisson berlaku equidispersi, sehingga dilakukan generalisasi terhadap distribusi Poisson menjadi distribusi COM-Poisson untuk menganalisis data diskrit yang equidispersi, overdispersi dan underdispersi. Generalisasi dari distribusi Binomial yang dapat menganalisis data dengan kejadian overdispersi dan underdispersi adalah distribusi COM-Poisson-Binomial. Pdf nya diperoleh dari distribusi COM-Poisson bersyarat dari penjumlahan dua distribusi COM-Poisson. Selain itu, dalam kajian ini juga dilakukan estimasi terhadap parameter-parameter dari COM-Poisson-Binomial dengan menggunakan Maximum Likelihood Estimation (MLE). Selanjutnya hasil estimasi ini dicoba pada data asosiasi sekunder dari kromosom di Brassika. M L E menghasilkan persamaan non-linier yang hasilnya digunakan untuk mencari nilai estimasi parameter θ dan parameter ν, persamaan non-linier tersebut diselesaikan dengan menggunakan metode Newton-Raphson. Dari proses tersebut didapatkan nilai estimasi parameter $\theta = 1.7421$ dan nilai estimasi parameter $\nu = 1.0328$

Kata Kunci: Distribusi Binomial; Distribusi COM-Poisson; Overdispersi, Underdispersi; Maximum Likelihood Estimation

1. Pendahuluan

Distribusi Binomial dan Poisson digunakan untuk menganalisis data diskrit. Namun, pada distribusi Poisson berlaku *equidispersi* (variansi dan mean bernilai sama). Sehingga distribusi Poisson tidak tepat digunakan untuk menganalisis data diskrit dengan kejadian *overdispersi* (nilai variansi lebih besar dari mean) atau kejadian *underdispersi* (nilai variansi lebih kecil dari mean). Untuk data diskrit dengan kejadian *overdispersi* dapat dimodelkan dengan distribusi Binomial Negatif yang merupakan gabungan distribusi Poisson dan Gamma, disamping itu untuk kasus *overdispersi* atau *underdispersi* yang disebabkan frekuensi nol terlalu banyak pada data dapat dimodelkan dengan *Zero Inflated Poisson (ZIP)* [1]. Distribusi untuk menganalisis data diskrit dengan kejadian *overdispersi* atau *underdispersi* adalah distribusi COM-Poisson (Conway-Maxwell-Poisson) [1]. Distribusi COM-Poisson merupakan generalisasi dari distribusi Poisson yang dikembangkan oleh Conway dan Maxwell pada tahun 1962.

Distribusi Binomial digeneralisasikan dengan berbagai cara. Dari segala generalisasinya ada beberapa generalisasi berbentuk perkalian dan pertambahan dari distribusi Binomial. *Probability density function* (pdf) dari distribusi perkalian Binomial adalah perkalian dari pdf dan faktornya. Itu membuat perbedaan variasinya lebih besar atau kurang dari variansi Binomial yang sesuai, hal itu bergantung pada nilai-nilai faktornya. Disisi lain distribusi Binomial yang bersifat pertambahan itu adalah campuran dari tiga model Binomial yang umum dan model korelasi Binomial yang mencakup variabel Bernoulli yang dependent [2].

Banyak literatur yang memperkenalkan generalisasi distribusi Binomial antara lain, literatur yang membahas tentang distribusi Binomial yang mempunyai 3 parameter yang merupakan generalisasi dari Binomial yaitu Beta-Binomial, dan korelasi distribusi Binomial [3]. Adapun literatur yang menjelaskan generalisasi lainnya dari distribusi Binomial yaitu distribusi COM-Poisson-Binomial [1], tetapi dalam literatur tersebut tidak membahas sifat-sifat dari distribusi COM-Poisson-Binomial.

Pada kajian ini membahas tentang sifat-sifat generalisasi dari distribusi Binomial yang bertype COM-Poisson

2. Tinjauan Pustaka

2.1. *Mean* dan *Variansi*

Mean atau nilai harapan dari peubah acak diskrit X dengan pdf $f(x)$ dapat didefinisikan sebagai [5]

$$\mu = E(X) = \sum_{x} x f(x) \tag{1}$$

Mean dinotasikan dengan μ atau $E(X)$.

Variansi dari peubah acak X yang dinotasikan dengan σ^2 atau $\mathrm{Var}(X)$ dapat didefinisikan sebagai [5]

$$\sigma^2 = \mathrm{Var}(X) = E[(X-\mu)^2] \tag{2}$$

dan standar deviasi dari X adalah $\sigma = \sqrt{\sigma^2}$.

2.2. Distribusi COM-Poisson

Distribusi COM-Poisson merupakan pengembangan dari distribusi Poisson yang ditemukan oleh Conway dan Maxwell. Distribusi COM-Poisson mampu memodelkan data yang mengalami *equidispersi, underdispersi* dan *overdispersi*.

Peubah acak X yang berdistribusi COM-Poisson dengan parameter $\lambda > 0$ dan $v > 0$ mempunyai pdf

$$f(x;\lambda,v) = \frac{\lambda^x}{(x!)^v}\frac{1}{Z(\lambda,v)}, \qquad x = 0,1,2,\dots\dots\dots \quad (3)$$

dimana,

$$Z(\lambda,v) = \sum_{x=0}^{\infty}\frac{\lambda^x}{(x!)^v} \approx \frac{e^{v\lambda^{\frac{1}{v}}}}{\lambda^{\frac{v-1}{2v}}(2\pi)^{\frac{v-1}{2}}\sqrt{v}} \quad [1]$$

dengan,

e : 2.71828
λ : parameter lokasi
v : parameter dispersi
$Z(\lambda,v)$: konstanta normalisasi

Subtitusi $Z(\lambda,v)$, Sehingga persamaan (3) menjadi

$$f(x;\lambda,v) = \frac{\lambda^x}{(x!)^v}\frac{1}{\frac{e^{v\lambda^{1/v}}}{\lambda^{\frac{v-1}{2v}}(2\pi)^{\frac{v-1}{2}}\sqrt{v}}} = \frac{\lambda^x}{(x!)^v}\frac{\lambda^{\frac{v-1}{2v}}(2\pi)^{\frac{v-1}{2}}\sqrt{v}}{e^{v\lambda^{1/v}}} = \frac{\lambda^{x+\frac{v-1}{2v}}(2\pi)^{\frac{v-1}{2}}\sqrt{v}}{(x!)^v e^{v\lambda^{\frac{1}{v}}}} \quad (4)$$

Beberapa sifat distribusi COM-Poisson adalah sebagai berikut
Mean : $\mu = E(X) = \lambda^{1/v} - \frac{v-1}{2v}$
Variansi : $\sigma^2 = \mathrm{Var}(X) = \frac{1}{v}\lambda^{1/v}$

2.3. *Maximum Likelihood Estimation*

Salah Salah satu metode yang dapat digunakan dalam mengestimasi parameter pada distribusi adalah *MLE*. *Maksimum likelihood* merupakan suatu cara yang mengarah pada sifat-sifat estimator yang diinginkan, terutama untuk sampel besar, yaitu dengan menggunakan nilai dalam parameter yang berhubungan dengan kemungkinan terbesar untuk data pengamatan sebagai perkiraan dari parameter yang tidak diketahui.

Untuk distribusi bersama dari n variabel $X_1,\dots\dots,X_n$ dengan nilai $x_1,\dots\dots,x_n$ dan $f(x_1,\dots\dots,x_n;\theta)$ adalah fungsi *likelihood*. Untuk $x_1,\dots\dots,x_n$yang tetap, fungsi *likelihood* nya adalah sebuah fungsi dari θ, ditulis dengan L(θ). Jika $X_1,\dots\dots,X_n$ adalah sampel acak dari $f(x;\theta)$, maka

$$L(\theta) = f(x_1;\theta)\dots f(x_n;\theta) \quad (5)$$

Dalam penerapan fungsi *likelihood*, $L(\theta)$ mempresentasikan distribusi bersama dari sampel acak, walaupun *maximum likelihood* juga dapat dipakai dalam kasus lain seperti dalam order statistik. Nilai θ yang memaksimalkan $L(\theta)$ juga akan memaksimalkan ln *likelihood* atau ditulis ln $L(\theta)$, untuk mendapatkan persamaan *maximum likelihood* yaitu [6]

$$\frac{d}{d\theta}\ln L(\theta) = 0 \quad (6)$$

2.4. Metode *Newton-Raphson*

Metode Newthon-Raphson adalah metode pendekatan untuk menyelesaikan persamaan nonlinear atau digunakan untuk menentukan titik saat fungsi maksimum. Titik pendekatan ke $t+1$ dituliskan sebagai [7]

$$\beta_{t+1} = \beta_t - [H_t]^{-1}G_t \quad (7)$$

dengan

β_{t+1} : vektor estimasi parameter pada iterasi ke $t+1$

β_t : vektor estimasi parameter pada iterasi ke t

$[H_t]^{-1}$: invers dari matriks Hessian yang isi dari matriks merupakan turunan kedua dari $\ln L(\beta)$.

G_t : vektor yang berisi turunan pertama dari $ln\, L(\beta)$.

Langkah-langkah metode Newton-Raphson:

1. Tentukan nilai awal β
2. Lakukan iterasi $\beta_{t+1} = \beta_t - [H_t]^{-1}G_t$ dimana $t = 0,1,2,3, \ldots.$
3. Iterasi berhenti jika salah satu kriteria dibawah ini terpenuhi:
   a. $|\beta_{t+1} - \beta_t| \leq \varepsilon_s$
   b. Banyaknya iterasi terlampaui

3. Hasil Penelitian dan Pembahasan

3.1.Kajian Distribusi COM-Poisson

Distribusi Distribusi COM-Poisson merupakan generalisasi dari distribusi Poisson. Dalam hal ini jika $v = 1$, maka pdf distribusi COM-Poisson merupakan pdf dari distribusi Poisson.

Untuk $v = 1$, subtitusi ke persamaan (4) diperoleh distribusi Poisson sebagai berikut:

$$f(x;\lambda,v) = \frac{\lambda^{x+\frac{v-1}{2v}}(2\pi)^{\frac{v-1}{2}}\sqrt{v}}{(x!)^v e^{v\lambda^{\frac{1}{v}}}}$$

$$f(x;\lambda,1) = \frac{\lambda^{x+\frac{1-1}{2v}}(2\pi)^{\frac{1-1}{2}}\sqrt{1}}{(x!)^1 e^{1\lambda^{\frac{1}{1}}}} = \frac{\lambda^{x+0}(2\pi)^0.1}{(x!)e^{\lambda}} = \frac{\lambda^x}{(x!)e^{\lambda}}$$

Sehingga diperoleh pdf distribusi Poisson $(x;\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}$

3.2.Pdf Distribusi COM-Poisson-Binomial

Distribusi Distribusi COM-Poisson Binomial dapat merepresentasikan *overdispersion* dan *underdispersion* relatif kepada distribusi Binomial pada umumnya. Ketika $v > 1$ mengalami *underdispersion* dan ketika $v < 1$ menunjukkan terjadinya *overdispersion* yang berhubungan dengan distribusi Binomial. Hal ini dikarenakan distribusi COM-Poisson-Binomial di definisikan dari distribusi COM-Poisson bersyarat yang merupakan penjumlahan dari dua variabel COM-Poisson yang independen.

Menggunakan definisi distribusi probabilitas binomial, dimisalkan $S = X + Y$ merupakan jumlah pengamatan. Sehingga untuk memperoleh $X = x$ dan $S = s$, maka perlu $X = x$ dan $Y = s - x$, maka diperoleh distribusi bersama dari X dan S

$$\begin{aligned} f_{X,S}(x,s) &= P[X = x, S = s] \\ &= P[X = x, Y = s - x] \\ &= f_{X,Y}(x, s - x) \end{aligned}$$

Karena variabel X dan Y merupakan variabel independen distribusi bersama dapat ditulis

$$
\begin{aligned}
f_{X,Y}(x,y) &= f(x)f(s-x) \\
&= \frac{\lambda_x^x}{(x!)^v Z(\lambda_x, v)} \frac{\lambda_y^{s-x}}{\left((s-x)!\right)^v Z(\lambda_y, v)} \\
&= \frac{(\lambda_x+\lambda_y)^s}{(s!)^v Z(\lambda_x, v) Z(\lambda_y, v)} \binom{s}{x}^v \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{s-x} \qquad (8)
\end{aligned}
$$

dengan distribusi marginal

$$
\begin{aligned}
f(s) = \sum_{x=0}^{s} f(x,s) &= \sum_{x=0}^{s} \frac{(\lambda_x+\lambda_y)^s}{(s!)^v Z(\lambda_x, v) Z(\lambda_y, v)} \binom{s}{x}^v \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{s-x} \\
&= \frac{(\lambda_x+\lambda_y)^s}{(s!)^v Z(\lambda_x, v) Z(\lambda_y, v)} \times \sum_{x=0}^{s} \binom{s}{x}^v \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{s-x} \qquad (9)
\end{aligned}
$$

Sehingga didapatkan distribusi bersyarat dari X diketahui $S = s$ dari distribusi bersama pada Persamaan (8) dan distribusi marginal pada Persamaan (9)

$$
f(x|s) = \frac{\binom{s}{x}^v \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{s-x}}{\sum_{x=0}^{s} \binom{s}{x}^v \left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x \left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{s-x}} \qquad (10)
$$

dengan $p = \frac{\lambda_x}{\lambda_x+\lambda_y}$ Persamaan (10) menjadi

$$
f(x|s) = \frac{\binom{s}{x}^v (p)^x (1-p)^{s-x}}{\sum_{x=0}^{s} \binom{s}{x}^v (p)^x (1-p)^{s-x}} \qquad (11)
$$

dari Persamaan (11) didapatkan pdf dari distribusi COM-Poisson-Binomial

$$
P(x; m, p, v) = \frac{\binom{m}{x}^v p^x (1-p)^{m-x}}{\sum_{x=0}^{m} \binom{m}{x}^v p^x (1-p)^{m-x}}, x = 0,1,.,m \qquad (12)
$$

dengan
m : jumlah pengamatan
p : peluang sukses

Distribusi COM-Poisson-Binomial merupakan generalisasi dari distribusi binomial. Hal ini diperoleh jika v=1 untuk m $\in \mathbb{Z}^+$, $p \in (0,1)$ dan $v \in \mathcal{R}$, pdf distribusi COM-Poisson-Binomial adalah pdf dari distribusi binomial.

$$
\begin{aligned}
f(x; m, p, 1) &= \frac{\binom{m}{x}^1 p^x (1-p)^{m-x}}{\sum_{x=0}^{m} \binom{m}{x}^1 p^x (1-p)^{m-x}} \\
&= \frac{\binom{m}{x} p^x (1-p)^{m-x}}{(1-p)^m + \binom{m}{1} p (1-p)^{m-1} + \cdots + p^m} \\
&= \binom{m}{x} p^x (1-p)^{m-x} \\
&= f(x; m, p)
\end{aligned}
$$

Teorema 2 [1]
Misal X adalah variabel COM-Poisson-Binomial dengan parameter $m, v, dan\ p$. Jika $m \to \infty$ dan $p \to 0$ dengan $\lambda = m^v p$ konstan, untuk $v \geq 0$ maka

$$
\lim_{m\to\infty} \frac{\binom{m}{x}^v p^x (1-p)^{m-x}}{\sum_{x=0}^{m} \binom{m}{x}^v p^x (1-p)^{m-x}} = \frac{\lambda^x}{(x!)^v} \frac{1}{\sum_{x=0}^{\infty} \frac{\lambda^k}{(x!)^v}} \qquad (13)
$$

Bukti

$$\lim_{m\to\infty} P[X=x|m,p,v] = \lim_{m\to\infty} \frac{\binom{m}{x}^v p^x(1-p)^{m-x}}{\sum_{x=0}^{m}\binom{m}{k}^v p^x(1-p)^{m-x}}$$

$$= \lim_{m\to\infty} \frac{\frac{\lambda^x}{(x!)^v}\frac{((m-x+1)\ldots\ldots m)^v}{(m^v-\lambda)^x}}{\sum_{x=0}^{m}\frac{\lambda^x}{(x!)^v}\frac{((m-x+1)\ldots\ldots m)^v}{(m^v-\lambda)^x}}$$

$$= \frac{\frac{\lambda^x}{(x!)^v}\lim_{m\to\infty}\frac{m^{xv}}{(m^v-\lambda)^x}}{\sum_{x=0}^{m}\frac{\lambda^x}{(x!)^v}\lim_{m\to\infty}\frac{m^{xv}}{(m^v-\lambda)^x}}$$

$$= \frac{\lambda^x}{(x!)^v}\frac{1}{\sum_{x=0}^{\infty}\frac{\lambda^x}{(x!)^v}} \qquad \blacksquare$$

Untuk mendapatkan struktur tipe COM-Poisson untuk pdf distribusi COM-Poisson-Binomial, Persamaan (12) dibagi pembilang dan penyebutnya dengan $(1-p)^m(m!)^v$.

$$P[X=x;m,p,v] = \frac{\binom{m}{x}^v p^x(1-p)^{m-x}}{\sum_{x=0}^{m}\binom{m}{x}^v p^x(1-p)^{m-x}} = \frac{\frac{1}{(x!\,(m-x)!)^v}p^x(1-p)^{-x}}{\sum_{x=0}^{m}\frac{1}{(x!\,(m-x)!)^v}p^x(1-p)^{-x}} \tag{14}$$

Selanjutnya, $\theta = \frac{p}{1-p}$ Disubtitusikan ke Persamaan (14) sehingga diperoleh pdf dari distribusi COM-Poisson-Binomial, yang dinotasikan sebagai X~CMPB(m,θ,v) menjadi

$$P[X=x;m,\theta,v] = \frac{1}{Z(\theta,v)}\frac{\theta^x}{(x!\,(m-x)!)^v}, x=0,1,,m \tag{15}$$

dengan

$Z(\theta,v) = \sum_{x=0}^{m}\frac{\theta^x}{(x!(m-x)!)^v}$

m : jumlah pengamatan

θ : parameter lokasi

v : parameter dispersi

$Z(\theta,v)$: konstanta normalisasi

3.3.Mean dan Varians distribusi COM-Poisson-Binomial

Pada bagian ini akan diuraikan cara memperoleh mean dan varians dari peubah acak X yang berdistribusi COM-Poisson-Binomial.

Dengan menggunakan Persamaan (1) didapatkan mean dari distribusi COM-Poisson-Binomial

$$E[x] = \sum_{x=0}^{m} x\,f(x)$$

$$= \sum_{x=0}^{m} x\frac{1}{Z(\theta,v)}\frac{\theta^x}{(x!\,(m-x)!)^v}$$

$$= \frac{1}{Z(\theta,v)}\sum_{x=0}^{m} x\frac{\theta^x}{(x!\,(m-x)!)^v}$$

$$= \theta\frac{1}{Z(\theta,v)}\sum_{x=0}^{m} x\frac{\theta^{x-1}}{(x!\,(m-x)!)^v}$$

$$= \theta . \frac{\partial \ln(Z(\theta, v))}{\partial \theta}$$

Selanjutnya, dengan menggunakan Persamaan (2) didapatkan varians dari distribusi COM-Poisson-Binomial

$$Var(x) = E[(X-\mu)^2] = \sum_{x=0}^{m} x^2 \frac{1}{Z(\theta,v)} \frac{\theta^x}{(x!\,(m-x)!)^v} - \left(\sum_{x=0}^{m} x \frac{1}{Z(\theta,v)} \frac{\theta^x}{(x!\,(m-x)!)^v}\right)^2$$

$$= \theta \left(\frac{\left(\sum_{x=0}^{m} x^2 \frac{\theta^{x-1}}{(x!\,(m-x)!)^v}\right)\left(\sum_{x=0}^{m} \frac{\theta^x}{(x!\,(m-x)!)^v}\right)}{\left(\sum_{x=0}^{m} \frac{\theta^x}{(x!\,(m-x)!)^v}\right)^2} \right) - \theta \left(\frac{\left(\sum_{x=0}^{m} x \frac{\theta^x}{(x!\,(m-x)!)^v}\right)\left(\sum_{x=0}^{m} x \frac{\theta^{x-1}}{(x!\,(m-x)!)^v}\right)}{\left(\sum_{x=0}^{m} \frac{\theta^x}{(x!\,(m-x)!)^v}\right)^2} \right)$$

$$= \theta \frac{\partial E[X]}{\partial \theta}$$

3.4.Estimasi Parameter dari Distribusi COM-Poisson-Binomial

Estimasi dari parameter distribusi COM-Poisson-Binomial bertujuan untuk mendapatkan estimator dari COM-Poisson-Binomial. Pada bagian ini akan dibahas mengenai bagaimana mencari estimator distribusi COM-Poisson-Binomial dengan menggunakan *Maximum Likelihood Estimation.* Dimisalkan $x_1, x_2, \ldots, x_n$ adalah sampel acak berukuran n yang berdistribusi COM-Poisson-Binomial. Sehingga, fungsi *likelihood* dari distribusi COM-Poisson-Binomial didefinisikan oleh

$$L(x_{1,\ldots\ldots,}x_n;\, \theta, v) = \left(\frac{\theta^{x_1}}{(x_i!\,(m-x_i)!)^v} \cdot \frac{1}{Z(\theta,v)}\right) \cdot\cdot \left(\frac{\theta^{x_n}}{(x_i!\,(m-x_i)!)^v} \cdot \frac{1}{Z(\theta,v)}\right)$$

$$= \left(\frac{1}{Z(\theta,v)}\right)^n \cdot \frac{\theta^{\sum_{i=1}^{n} x_i}}{\sum_{i=1}^{n} (x_i!\,(m-x_i)!)^v} \tag{16}$$

Oleh karena itu, diperoleh fungsi log *likelihood* sebagai berikut:

$$\ln L(x_{1,\ldots\ldots,}x_n;\, \theta, v) = \ln\left(\left(\frac{1}{Z(\theta,v)}\right)^n \cdot \frac{\theta^{\sum_{i=1}^{n} x_i}}{\sum_{i=1}^{n} (x_i!\,(m-x_i)!)^v}\right)$$

$$= \ln \theta \,.\, \sum_{i=1}^{n} x_i - v.\ln \sum_{i=1}^{n} (x_i!\,(m-x_i)!) - n.\ln Z(\theta, v) \tag{17}$$

Selanjutnya dicari turunan fungsi log *likelihood* terhadap masing-masing parameternya.

Untuk mendapatkan turunan pertama dari fungsi log *likelihood* terhadap θ dilakukan proses sebagai berikut :

$$\frac{\partial \ln L(\theta, v)}{\partial \theta} = \frac{\partial}{\partial \theta}\left(\ln \theta \,.\, \sum_{i=1}^{n} x_i - v.\ln \sum_{i=1}^{n} (x_i!\,(m-x_i)!) - n.\ln Z(\theta, v)\right)$$

$$= \left(\frac{\sum_{i=1}^{n} x_i}{\theta}\right) - \left(\frac{n\left(\sum_{x=0}^{m} x \cdot \frac{\theta^{x-1}}{(x!\,(m-x)!)^v}\right)}{\sum_{x=0}^{m} \frac{\theta^x}{(x!\,(m-x)!)^v}}\right) \tag{18}$$

Turunan pertama log likelihood terhadap v, dilakukan proses sebagai berikut :

$$\frac{\partial \ln L(\theta,v)}{\partial v}=\frac{\partial}{\partial v}\left(\ln\theta.\sum_{i=1}^{n}x_i-v.\sum_{i=1}^{n}\ln(x_i!\,(m-x_i)!)-n.\ln Z(\theta,v)\right)$$

$$=\left(-\sum_{i=1}^{n}\ln(x_i!\,(m-x_i)!)\right)+\frac{n\left(\sum_{x=0}^{m}\frac{\theta^x\ln(x!\,(x-m)!)}{(x!\,(m-x)!)^v}\right)}{\sum_{x=0}^{m}\frac{\theta^x}{(x!\,(m-x)!)^v}} \tag{19}$$

Persamaan (18) dan (19) selanjutnya digunakan untuk mencari nilai estimasi dari parameter θ dan v. Karena persamaan yang diperoleh adalah persamaan non linier Sehingga, untuk menyelesaikannya harus menggunakan metode numerik. Metode numerik yang digunakan adalah metode Newton Raphson. Penggunaan metode Newton Raphson dapat dilakukan melalui persamaan (7).

Adapun elemen-elemen yang terdapat dalam matriks hessian merupakan turunan kedua fungsi log *likelihood* adalah

$$H=\begin{bmatrix}\frac{\partial^2\ln L(\theta,v)}{\partial\theta^2} & \frac{\partial^2\ln L(\theta,v)}{\partial\theta\partial v}\\ \frac{\partial^2\ln L(\theta,v)}{\partial v\partial\theta} & \frac{\partial^2\ln L(\theta,v)}{\partial v^2}\end{bmatrix}$$

Dan G adalah vektor yang elemennya berisi turunan pertama fungsi log *likelihood*

$$G=\begin{bmatrix}\frac{\partial\ln L(\theta,v)}{\partial\theta}\\ \frac{\partial\ln L(\theta,v)}{\partial v}\end{bmatrix}$$

3.5.Penerapan Estimasi Parameter Distribusi COM-Poisson- Binomial

Data dalam contoh ini merujuk pada 337 pengamatan pada asosiasi sekunder dari kromosom di Brassika [8] yang disajikan di dalam Tabel 1, merupakan kejadian *underdispersi* dengan mean lebih besar daripada variansnya. Karena data dalam Tabel 1 *underdispersi* dan jumlah pasang bivalennya merupakan data cacah, sehingga dimodelkan dengan distribusi COM-Poisson-Binomial.

Tabel 1. Data asosiasi sekunder dari kromosom di Brassika

Jumlah pasang bivalen (x)	Jumlah data yang diamati	Peluang (x)	Peluang (x) berdistribusi CMPB (3, 1.7421, 1.0328)
0	32	0.0949	0.0473
1	103	0.3056	0.2562
2	122	0.3620	0.4464
3	80	0.2374	0.2501
Total	337	0.9999	1

Dengan menggunakan Persamaan (18), dengan nilai awal $\theta_0 = \bar{x}$,$v_0 = 1$ dan $\varepsilon = 0.0003$ hasilnya , didapatkan nilai estimasi parameter pada iterasi ke 11 yaitu $\theta = 1.7421$ dan $v = 1.0328$, sehingga data jumlah pasang bivalen (x) dalam Tabel 1 berdistribusi CMPB(m, θ, v) atau dapat ditulis CMPB (3, 1.7421, 1.0328) dengan pdf

$$f(x; m, \theta, v) = \frac{\theta^{\,x}}{(x!\,(m-x)!)^v} \frac{1}{\sum_{x=0}^{m} \frac{\theta^{\,x}}{(x!\,(m-x)!)^v}}$$

$$f(x; 3{,}1.7421\,{,}1.0328) = \frac{1.7421}{(x!\,(3-x)!)^{1.0328}} \frac{1}{\sum_{x=0}^{3} \frac{1.7421^{\,x}}{(x!\,(3-x)!)^{1.0328}}}$$

$x = 0{,}1{,}2{,}3$ menyatakan jumlah pasang bivalen. Dari pdf tersebut didapatkan probabilitas untuk masing-masing nilai x seperti pada Tabel 1 kolom ke 4.

Dari Tabel 1 didapatkan 2 pasang bivalen memiliki probabilitas tertinggi, sehingga modus data jumlah pasang bivalen tersebut sama dengan modus data berdistribusi CMPB (3, 1.7421, 1.0328). Sedangkan mean dari data berdistribusi CMPB (3, 1.7421, 1.0328) adalah

$$E(X) = \theta . \frac{\partial \ln(Z(\theta, v))}{\partial \theta}$$

$$E(X) = \theta \frac{\sum_{x=0}^{m} x \frac{\theta^{x-1}}{(x!\,(m-x)!)^v}}{\sum_{x=0}^{m} \frac{\theta^{x}}{(x!\,(m-x)!)^v}}$$

$$= (1.7421\,) \frac{\sum_{x=0}^{3} x \frac{1.7421^{\,x-1}}{(x!\,(3-x)!)^{1.0328}}}{\sum_{x=0}^{3} \frac{1.7421^{\,x}}{(x!\,(3-x)!)^{1.0328}}}$$

$$= (1.7421\;) \frac{3.6225}{3.3229}$$

$$= 1.8992$$

Mean tersebut hampir sama dengan mean dari data jumlah pasang bivalen (x) yaitu 1.7418. karena modus dan mean dari data jumlah pasang bivalen (x) hampir sama dengan modus dan mean data berdistribusi CMPB (3, 1.7421, 1.0328) maka nilai estimasi parameter yang diperoleh cocok dengan data.

4. Kesimpulan

Berdasarkan keseluruhan hasil penelitian dan pembahasan diperoleh kesimpulan sebagai berikut:

1. Bentuk generalisasi distribusi Binomial yang bertipe distribusi COM-Poisson menjadi COM-Poisson-Binomial dengan menjumlahkan dua distribusi COM-Poisson yang independen menghasilkan pdf distribusi COM-Poisson-Binomial $f(x; m, p, v) = \frac{\binom{m}{x}^v p^x (1-p)^{m-x}}{\sum_{x=0}^{m} \binom{m}{x}^v p^x (1-p)^{m-x}}, x = 0,1,.,m.$ Sehingga pdf COM-Poisson-Binomial yang bertipe COM-Poisson menjadi X~CMPB$(m, \theta, v) = \frac{1}{Z(\theta,v)} \frac{\theta^x}{(x!(m-x)!)^v}, x = 0,1, \ldots, m.$ Diperoleh juga sifat-sifatnya yaitu $E[x] = \theta . \frac{\partial \ln(Z(\theta,v))}{\partial \theta}$ dan $Var(x) = \theta \frac{\partial E[X]}{\partial \theta}$

   Dari hasil estimasi parameter dari distribusi COM-Poisson-Binomial dengan menggunakan *Maximum Likelihood Estimation* didapatkan parameter-parameter dari distribusi COM-Poisson-Binomial yang berupa persamaan-persamaan non linier sebagai berikut:

a. Untuk nilai parameter θ didapatkan persamaan non linier sebagai berikut:

$$\left(\frac{\sum_{i=1}^{n} x_i}{\theta}\right)-\left(\frac{n\left(\sum_{x=0}^{m} x \cdot \frac{\theta^{x-1}}{(x!\,(m-x)!)^{v}}\right)}{\sum_{x=0}^{m} \frac{\theta^{x}}{(x!\,(m-x)!)^{v}}}\right)=0$$

b. Untuk nilai parameter v didapatkan persamaan non linier sebagai berikut:

$$\left(-\sum_{i=1}^{n} \ln(x_i!\,(m-x_i)!)\right)+\frac{n\left(\sum_{x=0}^{m} \frac{\theta^{x} \ln(x!\,(x-m)!)}{(x!\,(m-x)!)^{v}}\right)}{\sum_{x=0}^{m} \frac{\theta^{x}}{(x!\,(m-x)!)^{v}}}=0$$

Untuk menyelesaikan persamaan non linier tersebut digunakan metode Newton-Raphson. Penerapan data asosiasi sekunder dari kromoson di Brassika pada *Maximum Likelihood Estimation* COM-Poisson-Binomial didapatkan nilai estimasi parameter $\theta = 1.7421$ dan $v = 1.0328$.

5. Daftar Pustaka

[1] Shmueli, G., T. P. Minka, J. B. Kadane, Borle, and P. Boatwright. 2005. “A Useful Ditribution for Fitting Discrete Data: Revival of The Conway-Maxwell-Poisson Distribution”. **Applied Statistics, Journal of Royal Statistical Society** 54 no. 1, hal.127-142.

[2] Altham, P.M.E. 1978. “Two generalizations of the binomial distribution”. **J. Roy. Statist. Soc. Ser**. C 27, hal. 162-167.

[3] Kupper, L.L., Haseman, J.K. 1978. “The use of a correlated binomial model for the analysis of certain toxicological experiments”. **Biometrics** 34, hal. 69-76.

[4] Borges, P., Rodrigues, J., Balakrishnan, N., and B. Jorge. 2014. “A COM-Poisson Type Generalization of the Binomial Distribution and its Properties and Applications”. **Statistics and Probability Letters** 87, hal.158-166.

[5] Walpole, R.E. **“Pengantar Statistika Edisi ke-3”**. Jakarta: PT. Gramedia Pustaka Utama.

[6] Bain, L.J., and Engelhardt, Max. 1991. **“Introduction to Probability and Mathematical Statistics 2nd edition”.** Belmont, California: Duxbury Press.

[7] Agresti, A. 2002. **“Categorical Data Analysis 2nd edition”.** John Wiley and Sons, New York.

[8] Skellam, J.G. 1948. “A probablility distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials”. **J.Roy. Statist.Soc.Ser.** B 10, hal.257-261.